Elder-proxy agreement concerning the elder's functional status and medical history: The impact of caregiver burden and depressive symptomatology

By: Kristie Long, S. Sudha, and Elizabeth J. Mutran

This is the peer reviewed version of the following article:

Schatzberg, K., Sudha S., & Mutran, E. J. (1998). Elder-proxy agreement concerning the elder's functional status and medical history: The impact of caregiver burden and depressive symptomatology. *Journal of the American Geriatrics Society, 46*, 1103-11.

which has been published in final form at https://doi.org/10.1111/j.1532-5415.1998.tb06648.x. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Abstract:

OBJECTIVES: To examine the influence of caregiver burden and depressive symptomatology on elder-proxy response concordance regarding the older person's functional status and medical history. **DESIGN:** Cross-sectional study via telephone interviews. **SETTING**: Community-dwelling older people and caregivers in North Carolina. **PARTICIPANTS:** 340 matched pairs of frail persons aged 65 and older and their respective caregivers.
MEASUREMENTS: Multidimensional Functional Assessment: The OARS methodology
RESULTS: Percent agreement on the ADL items ranged from 97.6% on personal hygiene to 99.7% for toileting, with moderate kappa coefficients. IADL percent agreement ranged from 71.5 to 93.7%, with fair to moderate kappa coefficients. Agreement among the medical history items ranged from 76.3 to 98.5% (κ = .138–.831). Response bias for the IADL composite measure is influenced marginally by caregiver burden (F[259] = 1.751. *P* = .098). Five of the individual IADL bias items are influenced significantly by burden. such that an increase in burden results in a greater likelihood that the caregiver will overstate disability compared with the rating by the older person. Response bias on the ADL scale was increased among persons who experienced more caregiver burden (OR = 1.096, 95% CI = 1.000, 1.192) and those who spent more hours providing care (OR = 1.012, 95% CI = 1.001, 1.024). Additionally, black caregivers were more likely than white caregivers to disagree with the older people on the ADL scale (OR = 2.73, 95% CI = 1.642, 3.809). A composite of the medical history items is influenced by the relationship of the caregiver to the older person; bias is more likely among adult children ((F[227] = 1.56, P = .081). **CONCLUSION**: Elder-proxy concordance is highest among ADL items, followed by medical history items and IADL items. Caregiver depressive symptomotology had no significant impact on elder-proxy response concordance on any of the three outcomes of interest: IADL and medical history bias and ADL disagreement. However, caregiver burden was marginally predictive of bias on the total ADL and IADL scales. Additionally, increased burden was significantly predictive of bias on five of the seven

individual items of the IADL scale, suggesting that the more burden a caregiver feels, the greater likelihood that s/he will overstate the older person's disability compared with self-report. These findings suggest that clinicians and researchers who use proxy reports to determine treatment regimens and complete data collection efforts may do so with confidence on ADL individual items and medical history items when the older person's frailty is marginal. However, caregiver burden may result in misleading representation of the older person's functional status, specifically in regard to IADL items.

Keywords: disability | elder care | caregiver burden

Article:

As the population of the country ages, and the incidence of physical and mental frailty increases, clinicians and researchers must frequently re-examine the ways in which information regarding older persons' functional status and medical history is gathered. The difficulty often faced by healthcare professionals and researchers when working with an aging population is the need to rely on surrogate or proxy responses when it is not possible to interview the older person directly. For the researcher, this situation may occur in up to 20% of cases among older community respondents or more than 50% of cases among hospitalized older people.[1, 2] The clinician may also rely on proxy reports of the older person's functional status and medical history when the patient is cognitively unable to provide reliable estimates him/herself.

The proxy's responses fill the gaps in data about the older person, substantially increasing the sample size for studies and forming the basis for treatment decisions by the healthcare professional. Proxy responses thus need to be scrutinized for how closely they match the response of the older subject on the same issue. This need for scrutiny forms the basis of a growing research trend examining concordance of the older person's and proxy's responses regarding the older person's health conditions and functional status.[1-5]

This research trend is still in its infancy, with many issues still unexamined. Most previous studies have examined patient-proxy response agreement among specific samples of patients, for example, those who have medical conditions such as Alzheimer's disease,[6-9] epilepsy,[10] coronary heart disease,[11] or hip fracture.[12] Others focus on care settings such as clinics or hospitals.[5, 6] The intensive care received by these subpopulations may limit the generalizability of the findings to a wider group of community-dwelling older people who may not be characterized by specialized needs, although proxy information is widely used in analyses of their situation.[1] Few studies examine elder-proxy response concordance among community-dwelling older people or a more representative group,[1, 9, 12-14] a gap we address by using a sample of community-dwelling, frail, black and white older people in North Carolina.

We examine first the degree of similarity between the older person's and the caregiver's responses to questions measuring the older person's activities and instrumental activities of daily living (ADLs and IADLs) and history of health conditions. Second, we extend the inquiry pioneered by Magaziner and colleagues,[1, 2] who have shown that characteristics of respondents and the caregiving context influence response bias in regard to ADLs and IADLs. Instead of considering one characteristic at a time as previous research has done, we consider a multivariate

model to predict response bias. We consider specifically the impact of two caregiver characteristics, caregiver burden and depressive symptoms, while controlling for demographic and socioeconomic characteristics of the proxy and the caregiving context (the familial relationship, caregiver's subjective health, and hours spent in caregiving). Earlier research investigated the impact of the older person's psychological impairment or depressive symptoms on response concordance,[3, 6] but only one study has examined the impact of proxy mood on concordance of rating depressive symptoms of their memory-impaired care recipient.[15] Scholars now call for specific examination of the way caregiver burden, physical health, and cognitive status of the proxy influence agreement.[6]

MATERIALS AND METHODS

Sample

We used survey data from a 1994 North Carolina study of frail older people and their caregivers that focused on ethnic differences in the use of adult care homes. Frailty was defined as requiring help in one or more ADLs or IADLs. Older people were selected from 29 North Carolina counties that met the following criteria: minority population greater than 10%, population size greater than 50,000, and at least five Family Care Homes and two Adult Care Homes per county. The 29 counties were then stratified into four groups by total population and minority representation, and 20 counties were selected randomly: urban, low minority (4), urban high minority (4), rural, low minority (6), and rural, high minority (6). An urban county was defined as one with a population size greater than 100,000, and a high-minority county was one with minority presence greater than 19%. Two samples of frail older people were selected from these 20 counties: one sample resided in the community (i.e., not in any type of adult care facility) and the other resided in adult care homes.

The sample of older people living in the community is used for this analysis. A random sample of persons aged 65 and over was drawn from each county using the Health Care Finance Administration's (HCFA) Medicare Beneficiary Utilization tape. A total of 4236 names were provided at the outset, of which 904 (21%) could not be located. Of the remaining 3332 persons, 411 (12%) refused to be screened for eligibility. The resulting 2921 older persons were screened for eligibility, of whom 71% were found to be ineligible for our study: they were not frail (57%); they lived in an institution, had a Department of Social Services address, or had moved out of the study area (10%); or they were deceased (4%). We randomly sampled 683 of the remaining 852 persons. Of this sample, 20% could not communicate and 2% refused to participate. Thus, a total of 537 older persons completed the telephone interview.

All caregivers of the original sample of 683 older adults were eligible to participate in the study regardless of whether the frail older person participated. The older people identified 563 family caregivers, and 507 (90%) completed the interview; 6.4% refused to participate and 3.5% could not be contacted or could not communicate. Because we are interested in the concordance of responses between caregivers and care recipients, we selected only matched pairs in which both the older person and his/her caregiver completed the interview (n = 361). Twenty-one cases were deleted from the analysis because of missing data, four cases because of missing responses to more than three ADL/IADLs, 12 cases because the caregivers did not respond to three or more of

the disease questions, and five cases because the caregiver did not complete the interview. Thus, 340 matched pairs were used for this analysis.

Measures

Dependent Variables: ADL/IADL Dependence

Caregivers and older adults were asked to rate independently their ability to perform seven ADLs and seven IADLs. The items were adapted from the OARS instrument to assess a persons' ability to perform tasks of daily living.[16] The older persons were asked if they could perform each task: (1) without difficulty, (2) with some difficulty, (3) unable to do it. The caregivers, however, were asked if the older person could perform each task (1) without help, (2) with some help, or (3) unable to do it. Because the wording of the response categories was slightly different for older person and caregivers, we opted to dichotomize the responses to reflect inability to carry out a task unaided, versus anything else. The complete dependence category ("unable to do it") was retained and coded as "1". We collapsed the categories "without help/difficulty" and "with some help/difficulty" into one category representing independence or limited assistance required, coded "0".

ADL tasks included dressing, eating, bathing, transfer out of bed, walking, toileting, and personal hygiene ($\alpha = .7387$). IADL questions were transportation, preparation of meals, grocery shopping, housework, laundry, finances, and medication administration ($\alpha = .8316$). The analyses of agreement and of bias are based on the dichotomy of each of the items and a composite measure of ADL and IADL bias.

To create the bias measures for individual item of Instrumental Activities of Daily Living, we calculated the difference between the elderly person's rating and the caregiver's (ep -eg) for each IADL dichotomy:

Positive Bias:

Elder rates more dificult than caregiver

$$1_{(ep)} - 0_{(cg)} = 1$$

Negative Bias:

Caregiver rates more dificult than elder

$$0_{(ep)} - 1_{(cg)} = -1$$

No Bias:

Elder and caregiver rate equally

$$1_{(ep)} - 1_{(cg)} = 0$$

or

$$0_{(ep)} - 0_{(ep)} = 0$$

The range was − 1 to 1 on each of the items; − 1 equals greater disability rated by the caregiver, zero represents perfect agreement, and +1 indicates greater disability reported by the older person. The composite measure of IADL bias included the bias scores on the individual IADL items entered simultaneously as the dependent variable in a multivariate regression model.

Although our original intent was to construct the dependent variables (i.e., individual item bias and composite measure of bias) for ADLs in the same manner as the IADLs, the high concordance between the older adult and caregiver responses on the ADL items did not create the acceptable distribution of bias that was necessary for our intended analysis strategy. Therefore, we used the same strategy illustrated for the IADL individual items and obtained a bias score for each ADL item. We then created a composite measure of ADL bias by summing the bias across all ADL items (range −6 to +2). We dichotomized the dependent variable to represent agreement and disagreement: "0" if there was no bias (n = 310) and "1" if there was bias in either direction for the composite ADL scale (n = 30). We were concerned that this strategy might misrepresent bias because the elder-caregiver pair may score a for bathing and a "+ for toileting, leading to a composite score of "0." To address this concern, we examined each pair to determine that we were not creating a false dependent variable by canceling out positive and negative bias on different items. Because bias was so minimal among the ADL items, our constructed dependent variable created an acceptable measure of response concordance. Thus, we used the composite measure as the only dependent variable of ADL bias.

Medical Conditions

Older people and caregivers were asked if the care recipient had ever seen a doctor for a prescribed list of medical conditions: hypertension, Alzheimer's disease, stroke or cerebrovascular accident, lung disease (asthma, bronchitis, or emphysema), arthritis or rheumatism, diabetes, cancer, depression, a broken hip, osteoporosis, amputation, rheumatic heart disease, coronary heart disease, epilepsy, kidney disease, anemia, eye disorder (e.g., cataracts, glaucoma), or mental health condition other than depression. In order to get accurate responses, the interviewers also provided alternative local phrasing for some diseases such as "sugar disease" for diabetes or "fragile/soft bones" for osteoporosis, and "high blood" for hypertension. Biases for the individual medical history items were constructed as follows:

Positive Bias:

Elder rates s/he has disease and caregiver does not

$$1_{(ep)} - 0_{(cg)} = 1$$

Negative Bias:

Caregiver rates elder has disease and elder does not

$$0_{(ep)} - 1_{(cg)} = 1$$

No Bias:

Elder and caregiver rate equally

$$1_{(ep)} - 1_{(cg)} = 0$$

or

$$0_{(ep)} - 0_{(ep)} = 0$$

Covariates: Demographic and Socioeconomic Status

The following information was ascertained for the caregivers: gender, ethnicity (black and white), age (calculated by year of birth), marital status, employment status, number of children living at home, annual household income (a scale ranging from 1 to 8), and education (number of school years completed). Information gathered about the older persons was similar and included gender, ethnicity, age, marital status, annual household income, whether they received Supplemental Security Income (SSI) and Medicaid, and perceived health. In this paper, only caregiver characteristics are considered as explanatory variables in the multivariate model.

Caregiving Situation

Indicators of the caregiving situation included the number of hours spent in the caregiving relationship (hours per week), the caregiver's subjective health, and the relationship to the care recipient. Caregiver health was measured by asking, “In general, would you say your health is (1) excellent, (2) very good, (3) good, (4) fair, (5) poor.” A higher score indicated a negative perception of the caregiver's health. The relationship variables were dummy variables indicating spouse or adult child, with the reference category as “other” caregiver.

Caregiver Burden

Caregiver burden was constructed from 10 questions.[17, 18] The following four questions were coded from 1 (not at all) to 4 (a lot): (1) to what extent does caring for the care recipient (CR) make your life difficult; (2) how much stress do you feel; (3) to what extent does helping the CR interfere with your personal life; (4) do you feel the situation is fair to you (reverse coded for consistent direction). Six additional questions were coded on a 3-point scale from *No*, (1) *Somewhat* (2) to *Yes* (3). The questions were: (1) not enough time for you to spend with your family, (2) need to take care of the care recipient when you were not feeling well enough, (3) conflicts with job demands, (4) problems that family members had with the care recipient, (5) lack of support from other family members, (6) lack of time for self. The high internal consistency of this measure is indicated by $\alpha = .8423$.

Depressive Symptomatology

Depressive symptomatology was measured by asking the following five questions: During the past month, how much of the time (1) have you been a nervous person, (2) have you felt calm and peaceful, (3) have you felt down-hearted and blue, (4) have you been a happy person, and (5) have you felt so down in the dumps that nothing could cheer you.[19] Responses ranged from 1 to 4 and were coded such that a higher score represented more depressive symptoms ($\alpha = .8408$).

Analysis

We compared responses by the older adults and their caregivers on the dichotomized ADL and IADL tasks and medical history items in terms of percent agreement, percent disagreement indicating direction of discrepancy (older person rates more difficulty than caregiver or caregiver rates more difficulty than older person) and Cohen's kappa statistic. The kappa statistic allowed us to capture the proportion of agreement beyond what is attributable to chance alone. These kappa statistics are unweighted because of the binary response format. The guidelines for interpreting the kappa statistic are as follows: 1 = perfect agreement; ≥ 0.81, almost perfect agreement; 0.6–0.8, substantial agreement; 0.4–0.6, moderate agreement; ≤ .4, slight to fair agreement.[20]

Next, we determined the extent to which caregiver characteristics influenced concordance between respondents after controlling for sociodemographic variables and the context of the caregiving situation. Our analytical strategy builds on previous research that examined one characteristic at a time to explain response bias by developing an empirical data-based model to explain agreement and the direction of bias in ADL, IADL, and medical history responses. Although no studies have examined caregiver characteristics on the extent to which they predict bias in reporting of functional status and medical history, a similar strategy was used to evaluate the relationship between demographic and mental health variables of the older person with degree of concordance between self-reported falls and chart-recorded falls.[21]

We conducted multivariate analysis of variance (MANOVA) with a two-tailed test of significance ($P < .05$) to determine how multiple covariates influence the overall model of IADLs and medical history items and the individual items that comprise these scales. When explanatory variables are marginally significant *(*$P < .10$), we report these results as well. MANOVA is used traditionally in an experimental design to test the effects of different treatments (between-subject factors) on multiple outcomes (within subject factors).[22, 23] In our study, the treatment effects are the covariates, and the multiple outcomes are the individual items of the IADL and medical history models. Although we do not intentionally manipulate our treatments (e.g., burden and depressive symptoms), they vary within our population and, thus, serve as treatment effects.

There are several strengths to this methodology. First, MANOVA allows us to examine the influence of multiple covariates on multiple outcomes simultaneously. Subsequently, we are able to address the extent to which caregiver burden and depressive symptoms result in response bias of the total IADL and medical history models after controlling for the sociodemographic characteristics of the caregiver and the caregiving situation. Second, by examining the parameter

estimates of covariates across the individual items that comprise the IADL and medical history scales, we determine the impact of the covariates on each individual item versus the composite measure (e.g., how does burden influence bias regarding shopping or hypertension). We can interpret the extent and direction of bias by examining the level of significance and sign of the parameter coefficient, respectively. The ability to examine how burden and depressive symptoms influence bias (e.g., does burden or depressive symptoms increase the likelihood that the caregiver will over- or understate IADL ability) is a significant contribution of the MANOVA strategy.

Although our intention was also to utilize the MANOVA strategy for the ADL scale, the high concordance of responses between the older person and the caregiver did not create an acceptable distribution of bias to use this technique. Therefore, we dichotomized our dependent variable: "0" if there was agreement and "1" if there was disagreement in either direction for the composite ADL scale. We conducted logistic regression analysis using the same covariates of the IADL and medical history models to examine the extent to which the covariates influence agreement or disagreement for the total ADL scale. Logistic regression was used because of the dichotomous, categorical outcome variables in the model. We were unable to examine the direction in which the covariates influence bias or how the covariates influence the individual items of the ADL bias scale because of the dichotomous nature of our dependent variable and our analytical strategy.

RESULTS

The caregivers in this sample were predominantly middle-aged, married women with household incomes between $15,001 and $25,000 per year. Approximately half were employed outside the home at the time of the study. As a result of the sample design, about half were black. A third had at least one child at home. More than half were adult children caring for their ailing parent and approximately one-fourth were spouses of the older person. They spent an average of 21 hours per week in the caregiving role (Table 1).

Table 1. Demographic Background of Caregivers

	N = 340
Gender(%)	
female	75.3
Ethnicity (%)	
Black	50.3
White	49.7
Age, mean in years ± SD	56 ± 14.02
Marital status (%)	
Married	62.9
Not married	37.1
Work status (%)	
Employed full-time	40.3
Employed part-time	11.2
Not working	48.4
At least one child at home (%)	35.5
Education, mean years ± SD	12.3 ± 3.15

	N = 340
Income (%)	
$7500 or less	14.3
$7501-$15,000	23.6
$15,001-$25,000	25.6
$25,001-$35,000	13.0
$35,001-$50,000	14.0
$50,001-$75,000	6.3
$75,001-$100,000	2.3
$100,001 and over	1.0
Refused or don't know	11.5
Average income, mean ± SD	3.22 ± 1.63
Hours spent caregiving per week, mean ± SD	21.4 ± 30.8
Relationship to older person (%)	
Spouse	26.2
Child	52.4
Other	21.5
Perceived health, mean ± SD	2.68 ±1.1 (Range 1–5)
Burden, mean ± SD	15.5 ± 4.8 (Range 10–30)
Depressive symptoms, mean ± SD	8.61 ± 2.6 (Range 5–20)

The majority of the older people in our sample was women (69.4%) and about half was widowed (49.7%). They averaged 76 years of age (range 65–98) and perceived their health to be fair. Approximately half of our sample was black. Forty percent of our sample reported a household income of less than $7500, and 71% reported income of less than $15,000. Eighteen percent received Supplemental Security Income and 33% received Medicaid.

The percent agreement among the ADLs ranged from 97.6 to 99.7% (Table 2). The kappa statistics of the ADL items found the agreement between the raters moderate on four of the ADLs, with the remaining three ranging from .166 for bathing to .856 for toileting.

Table 2. Caregiver-Frail Elder Response Agreement Regarding Functional Status

Items	n	Caregiver Rated more Difficulty (%)	Elder Rated More Difficulty (%)	Agreement (%)	Kappa (Standard Error)
ADLs					
Dress/undress	340	2 (0.6)	3(1.5)	333 (97.9)	.452(0.17)
Eating	340	1 (0.3)	1 (0.3)	338 (99.4)	.497(0.31)
Bathing	340	2 (0.6)	16(4.7)	322 (94.7)	.166(0.17)
Transfer in/out of bed	340	2 (0.6)	0	338 (99.4)	.664 (0.22)
Walking (move around the house)	340	3 (0.9)	2 (0.6)	335 (98.5)	.437 (0.22)
Toileting	340	1 (0.3)	0	339 (99.7)	.856(0.14)
Personal hygiene	340	5(1.5)	3 (0.9)	332 (97.6)	.417(0.17)
Total score					.661 (0.15)*
IADLs					
Get to places out of walking distance	340	38 (11.2)	59 (17.4)	243 (71.5)	.263 (0.06)
Meal preparation	338	22 (6.5)	14(4.1)	302 (89.3)	.539 (0.07)
Shopping	336	24 (7.1)	31 (9.2)	281 (83.6)	.593 (0.05)
Housework	336	39(11.6)	32 (9.5)	265 (78.9)	.435 (0.06)

Items	n	Caregiver Rated more Difficulty (%)	Elder Rated More Difficulty (%)	Agreement (%)	Kappa (Standard Error)
Laundry	337	20 (5.9)	24 (7.1)	293 (86.9)	.565 (0.06)
Finances	332	25 (7.5)	17(5.1)	290 (85.3)	.391 (0.08)
Medication administration	335	13(3.9)	8 (2.4)	314(93.7)	.400(0.11)
Total score					.644(0.05)*

* Intraclass correlation coefficient of composite measures.

Agreement among the IADL responses ranged from 71.5 to 93.7%. The item with least concordance was “getting to places out of walking distance” most reliable was “medication administration.” The caregiver reported greater dependence than did the older person on four of the seven items. However, the caregiver was more likely to overstate the older person's ability to get to places out of walking distance, shop for groceries, and do laundry. The kappa values among the IADLs were moderate (.263–.593).

The agreement among the medical history items ranged from 78.0 to 98.5% (Table 3). Arthritis and amputation were in least and most agreement, respectively. The kappa statistics for the 17 medical history questions were as follows: almost perfect (2), substantial (4), moderate (5), and fair agreement (6). The lower kappa statistics were often associated with very high percent agreement because of the lack of response variation. This occurred for the items Alzheimer's disease, rheumatic heart disease, and epilepsy.

Table 3. Caregiver-Elder Response Agreement Regarding Disease History

Items	n	Disease (CG)/No Disease (Elder) (%)	No Disease (CG)/Disease (Elder) (%)	Agreement (%)	Kappa (Standard Error)
Hypertension	333	22 (6.6)	35 (10.5)	276 (82.9)	.603 (0.05)
Alzheimer's	329	3 (0.9)	2 (0.6)	324 (98.5)	.278 (0.23)
Stroke	339	25 (7.4)	16(4.7)	298 (87.9)	.629 (0.05)
Lung disease	338	27 (8.0)	30 (8.9)	281 (83.1)	.547 (0.05)
Arthritis	336	41 (12.2)	33 (9.8)	262 (78.0)	.461 (0.05)
Diabetes	337	9 (2.7)	14(4.2)	314(93.2)	.847 (0.03)
Cancer	338	24(7.1)	13(3.8)	301 (89.1)	.654 (0.05)
Depression	332	30 (9.0)	36(10.8)	266 (80.1)	.260 (0.07)
Broken hip	340	12(3.5)	2 (0.6)	326 (95.9)	.646 (0.09)
Osteoporosis	333	24 (7.2)	15(4.5)	294 (88.3)	.415(0.08)
Amputation	338	2 (0.6)	3 (0.9)	333 (98.5)	.831 (0.07)
Rheumatic heart disease (RHD)	333	4(1.2)	7 (2.1)	322 (96.7)	.138(0.14)
Coronary heart disease (CHD)	331	32 (9.7)	31 (9.4)	268 (81.0)	.576 (0.05)
Epilepsy	338	6(1.8)	5(1.5)	327 (96.7)	.336(0.15)
Kidney disease	334	24 (7.2)	41 (12.3)	269 (80.5)	.347 (0.06)
Anemia	332	29 (8.7)	25 (7.5)	278 (83.7)	.306 (0.07)
Eye disorders	337	46 (13.6)	34 (10.1)	257 (76.3)	.524 (0.05)

The next step of our analysis was to examine the impact of our explanatory variables on rater agreement or bias. The purpose was to examine the influence of caregiver burden and depressive

symptomotology on three outcome variables: IADL dependency bias, ADL agreement and disagreement, and medical history bias.

Each IADL bias item (e.g., bias in elder-caregiver response to the ability of the older person to do laundry) was entered simultaneously as seven dependent variables in a MANOVA model, along with the explanatory covariates (Table 4). Caregiver burden was a marginally significant predictor of bias on the total model (F[259] = 1.751, $P < .098$). Although the total model was useful in determining which factors influence bias on the composite IADL model, parameter estimates of the individual items comprising the IADL scale allowed us to examine the direction of bias in reporting and the significant predictors of bias on the individual items.

Table 4. . Multivariate Model of Caregiver Characteristics to Influence ADL*, IADL†, and Medical History Bias†

	ADLs OR (95% CI)	IADLs F-ratio, df = 259 (P value)	Medical History F-ratio, df = 227 (P value)
Gender	.546 (−0.478, 1.570)	.139 (.995)	1.438 (.125)
Ethnicity	2.726 (1.642, 3.809)	.469 (.856)	1.067 (.388)
Age	.987(0.944, 1.031)	.745 (.634)	1.146 (.314)
Marital status	.989 (−0.090, 2.069)	.601 (.755)	1.045 (.411)
Work status	.699 (−0.345, 1.742)	.601 (.755)	.721 (.772)
Children	1.156(0.653, 1.660)	.615(743)	1.033 (.423)
Income	.782(0.341, 1.225)	1.325 (.239)	.747 (.744)
Education	.958(0.794, 1.122)	1.605 (.134)	.681 (.811)
Caregiving hours	1.01 (1.001, 1.024)	1.407 (.203)	.842 (.637)
Adult child caregiver	.739 (−0.342, 1.821)	1.335 (.234)	1.559 (.081)
Spouse caregiver	.553 (−0.985, 2.091)	.593 (.761)	.972 (.489)
Caregiver health	.712(0.225,1.199)	.592 (.762)	1.264 (.222)
Burden	1.10(1.000, 1.192)	1.751 (.098)	1.190 (.277)
Depressive symptoms	.941 (0.733, 1.150)	.343 (.934)	.606 (.877)

* Odds ratio using logistic regression.
† Wilk's lambda test statistic using MANVA.

The caregiver reported significantly more disability than the older person on five of the seven IADL items when the caregiver expressed more burden: meal preparation, shopping, housework, laundry and managing finances (see Table 5).

Table 5. Influence of Caregiver Burden on the Individual Items of the IADL Scale Using MANOVA*

	Get to Places out of Walking Distance	Meal preparation	Shopping	Housework	Laundry	Finances	Medication Administration
Burden	−.413	−2.023	−2.127	−2.040	−2.405	−1.999	−1.670
	(.008)	(.005)	(.006)	(.007)	(.005)	(.005)	(.004)
	.680	.044	.034	.042	.017	.047	.096

* β, (SE), *p*

Additionally, medication administration was marginally significant in the same direction. There was a linear relationship between burden and IADL bias, indicating that the less burdened a caregiver feels, the more likely they are to understate disability compared with the older person.

Depression was not a significant predictor of bias on the IADL model or on any of the individual items that comprise the IADL scale.

To examine the impact of the explanatory variables on the ADL composite measure, we dichotomized bias of the composite measure into agreement (0) and disagreement (1) and conducted logistic regression. Table 4 illustrates that increased hours in the caregiving relationship (OR = 1.012, 95% CI = 1.001, 1.024) and more caregiver burden (OR = 1.096, 95% CI = 1.000, 1.192) resulted in greater disagreement. Additionally, black caregivers were almost three times more likely to disagree with the older person than white caregivers (OR = 2.73, 95% CI = 1.642, 3.809).

Finally, we conducted a MANOVA analysis with the medical history items as the dependent variables (Table 4). Relationship to the older person influenced bias marginally (F[227] = 1.56, *P* = .081); adult children were more likely to be in disagreement than were other caregivers. In examining each medical history item independently, adult child caregivers overstated the presence of hypertension in older people and understated a history of kidney disease. To ensure that our results were not influenced by lack of response variation on the individual items, we conducted the same analysis with only those medical history items with a combined bias greater than 10%. The only change was that gender became a marginally significant predictor of bias of the overall scale (F[240] = 1.658, *P* = .084) and a significant predictor of bias for hypertension and cancer (F[l] = 3.923, *P*= .049 and F[l] = 5.657, *P* = .018, respectively). These findings suggest that male caregivers are less likely than female care providers to report a history of hypertension and cancer of the older person in their care.

DISCUSSION

As the population ages, the need for informal caregivers will inevitably rise. Increasingly, family members or other caregivers will be called upon to provide care in order to postpone institutional care for the older persons for as long as possible. The supply of informal caregivers, however, will likely decrease with the lower fertility rates and longer life expectancy of current and future older adults. Nonetheless, they will be expected to assume the role of caregiver without regard for the positive or negative consequences of caregiving.

Caregivers of frail older people are vulnerable to depression, stress, increased role strain, and burden caused by the extraordinary demands that accompany care provision.[24-29] Caregiving literature, however, has only begun to address how the strains associated with care provision affect the caregiver's ability to assess an older person's health status or health needs. Nonetheless, informal helpers are often called on to measure and provide for older persons' health needs, including acting as a surrogate in the health care system on the older person's behalf. The health information provided by the caregiver may be used to make healthcare decisions in cases where the older person cannot answer for him/herself. Therefore, it becomes increasingly important to consider how caregiver characteristics, in general, and caregiver burden and depressive symptoms, in particular, affect agreement between their responses and those of the older persons in their care regarding health and functional status measures.

The aims of this study were twofold. First, we examined response concordance for a dichotomous response format of ADL/IADL status and medical history items. Second, and of primary interest, we investigated how a proxy's characteristics, specifically burden and depressive symptomatology, influenced response bias regarding the older person's functional and health status.

With respect to the first issue, we find that older people and their proxies agree most often on ADL items and objective medical history items. Our findings are consistent with a previous study by Kiyak and colleagues, who found more agreement on functional status measures when rated by healthy older people and their families compared with demented patients and their proxies.[9] Other studies have shown moderate to good agreement when investigating response concordance among functional status items and observable, concrete, comorbidities.[1, 2, 5] The low prevalence of ADL dependency among this community-dwelling sample and the dichotomization of responses likely influenced the high level of agreement.

The only exception to the high concordance among ADLs in our sample is bathing. Older people were more than four times more likely (18: 4) than caregivers to rate themselves completely dependent. This finding is especially disconcerting because the bathroom may be a likely place for falls. Slick floors, lack of appropriately placed grab bars, wet tubs, and the older person's continued desire for privacy create an optimal scenario for falls. Fleming and Pendergast found that 57% of the falls by adult care facility residents occurred in the residents' rooms, with the private or shared bathroom the most frequent locus.[30]

Among the medical history items, arthritis had the lowest agreement. We were not surprised at this finding because the symptoms of arthritis are often invisible, not discussed, and, consequently, open to subjective interpretation. Additionally, arthritis symptoms are often transient, which could result in a different rater response on any given day. Amputation, however, is a visible and unambiguous condition; thus agreement among raters is improved. The high rate of agreement with respect to Alzheimer's disease is likely attributable to our sample, in which very few people had the disease (0.9% prevalence).

There was greater disagreement between caregivers and older people on the IADL items than on the ADL scale and medical history questions. Caregivers reported that the older people had greater difficulty on four of the seven items: meal preparation, housework, finances, and medication administration. On the remaining three items, however, the older people stated that they were completely unable to get to places out of walking distance, do laundry, and go shopping, whereas the caregiver rated that they were able to do these activities with some assistance or without help. Magaziner and colleagues also found that the proxy responses for aged hip fracture patients were more likely to report that the older person was disabled in the areas of meal preparation, housework, finances, and medication administration. They found no difference between patient and proxy on walking or shopping, and doing laundry was not investigated.[2]

Because we do not have an independent functional status verification to validate the responses of the older people or their caregivers, we do not know whether the caregivers overstate disability or the older people over-report their abilities on these measures. The implications of these

findings echo previous research that calls for accurate and specific phrasing of questions regarding the older person's status in order to lower potential for disagreement caused by differences in judgment or observation.[31]

With respect to our investigation of caregiver characteristics on response bias, we find that burden has an impact on the response agreement among the individual items of the IADL scale. Although it was only marginally significant, caregiver burden was the most influential factor explaining bias in the IADL summary score and a marginally significant predictor of disagreement for the ADL composite measure. Burden did not impact response concordance on the medical history scale or the individual items significantly.

Because we do not have a "gold standard" to measure "true" IADL status, we acknowledge that older persons may be underreporting their level of disability, not that caregivers are over-reporting the older people's frailty. We also must acknowledge that measures of caregiver burden and depressive symptomotology are even more subjective and require similar interpretation. Nonetheless, when caregivers report no burden, bias is reversed, which represents the linear relationship between burden and IADL bias. This finding is noteworthy for future investigations of senior functional status and caregiver burden.

With regard to depressive symptomatology of the caregiver, we did not find that this composite measure influenced bias among any of our outcome measures. Our findings may be a result of the relatively few caregivers who experienced depressive symptoms. Furthermore, the scale utilized in this study does not measure clinical depression, which could have a greater impact on response agreement. Our results are supported by the findings of Teri and Truax, who found that caregiver mood did not interfere with the primary caregivers' ability to identify depression among their memory-impaired care recipients correctly.[15] The caregivers did not appear to be biased by their own mood in their assessment of the patients' mood, which led the authors to conclude that caregivers' reports of patient mood may be considered confidently. Our study provides similar support for their hypothesis with regard to functional status and medical history. We do not, however, have objective measures of our dependent measures or of depressive symptoms; thus, further studies replicating these results are needed.

We did not expect the sociodemographic factors to predict significant bias among any of our outcome measures. Unexpectedly, we found that black caregivers are almost three times more likely than white caregivers to disagree regarding ADL measures. To determine the direction of bias, we examined the summary measure of ADL bias before dichotomization. Most black and white caregivers are in agreement with the older person (91%). When they do not agree, however, black caregivers are more likely than whites to underreport the older person's disability. Ethnic differences regarding reporting have not previously been studied, and we had no reason to expect ethnicity to be a significant explanatory variable.

An additional important finding was the positive relationship between caregiving hours and disagreement. The more hours spent in the caregiving role, the more likely caregivers are to be in disagreement. This is contrary to expectations because previous studies have shown that caregivers who live with the older person show higher correlations of ADL composite measures than caregivers who do not live with the older person.[2] The findings are likely attributable to the

increased burden experienced by the caregivers when they spend more hours providing care to the older person. Similarly, the positive correlation between the hours caregiving and burden are likely influencing the positive outcome in the regression analysis.

As Magaziner et al.[1] argue, knowing that caregiver (proxy) characteristics influence response agreement and bias might be of more academic than practical interest because most researchers or practitioners do not have the luxury of choosing the "ideal" proxy. With respect to research, our results suggest that controlling for proxy characteristics to minimize the potential impact of caregiver burden would be an optimal strategy; data collection efforts should attempt to include information on caregiver characteristics when possible. In addition, greater attention should be given to how questions are worded, optimizing questions that examine concrete conditions.

The practice recommendations arising from these results include increased efforts at eliciting as much information as possible from a "gold standard" and/or independent professional verification.[1] With respect to medical conditions, documenting the presence of a range of conditions at the time of an annual physical examination would be a fairly simple modification to existing procedures when a complete patient history is not present. With respect to the more subjective areas of assessing functional limitations, one could consider incorporating ADL and IADL measures into the annual examination. The older person/ caregiver/ independent verification by a healthcare professional would be ideal. Updating this information each year would provide a more complete picture of the patient's medical history at present and over time. Additionally, our findings suggest that the healthcare providers assess the caregivers' current level of involvement with the older patients and their sense of caregiver burden when their reports are serving as proxies for older persons' functional status.

It is important that several caveats be considered when interpreting the results of this study. First, we dichotomized responses differently from previous studies. Most previous studies dichotomized responses contrasting "complete independence" with "some or total dependence".[1, 2, 5] We, however, chose to collapse the "independent" and "some dependence" items into one category because of the inconsistent wording used to gather the older persons' and caregivers' reports of functional status. Because "unable to do it" was asked identically for both respondents, we retained that category as our measure of complete dependence. By examining the reliability across the individual items of the functional status scale, we find that our percent agreement and kappa coefficients are similar to those reported previously, which provides some reassurance that our dichotomy is acceptable. Additionally, we conducted a subsequent analysis maintaining the three-response format for functional status measures and found that percent agreement continued to be higher for items that are more concrete and observable (ADLs vs IADLs), similar to our current investigation. The agreement and weighted kappa coefficients of the additional analysis were less favorable than our current study. These findings are likely attributable to the dichotomization of responses rather than to the wording change. In this subsequent analysis, we also found burden to be the primary (although marginal) predictor of bias among the IADL scale items and in the same direction among the individual items.

Second, it is important to note that our study is the first to utilize a multivariate approach to predict bias between caregivers and older people. Because this strategy has not been utilized previously, we must interpret our results with caution until they can be replicated. We also

recognize that there are multiple variables that could be included as predictors of bias, including care recipient characteristics. Future examination of elder-proxy response concordance should investigate other possible explanations of bias.

Third, our study makes a unique contribution to the inquiry into proxy-elder response concordance by utilizing a community-dwelling, ethnically diverse sample. Our results can be extended to a more general population of older adults, in contrast to studies that consider specific subpopulations based on setting of care or presence of a certain medical condition. We had no a priori reason to expect ethnic differences in our results, and the finding that there is a greater disagreement among blacks than whites warrants further investigation.

As people age and become frail, the involvement of family members or other caregivers becomes essential to the older person's well-being. In turn, caregiving can lead to an overwhelming sense of burden. As we have shown, burden may influence how s/he perceives the needs of the older person and, ultimately, affect agreement between the older person and the caregiver. The perceptions of disability could, in turn, influence care provision and the older person's health status. Thus, clinicians should be aware of the biases identified here and make efforts to include both the caregiver and the older person when conducting clinical assessments and gathering patient histories. Actual measures of functional status can then verify their reports. Additionally, researchers who choose to impute values of the functional status of an older person based on proxy reports should be aware of the biases that may exist as a result of caregiver characteristics and interpret their results with appropriate caution.

ACKNOWLEDGMENTS

The authors are grateful to Jay Magaziner at the University of Maryland-Baltimore for his guidance regarding this manuscript. We also would like to thank Gary Gaddy at the Institute for Research in Social Science at the University of North Carolina-Chapel Hill for his statistical consulting. Lastly, we acknowledge with gratitude the financial support from the National Institute on Nursing Research.

REFERENCES

1. Magaziner J., Bassett SS, Hebel JR, Gruber-Baldini A.. Use of proxies to measure health and functional status in epidemiologic studies of community-dwelling women aged 65 years and older. *Am J Epidemiol* 1996; **143**: 283– 292.
2. Magaziner J., Simonsick EM, Kashner TM, Hebel JR. Patient proxy response comparability on measures of patient health and functional status. *J Clin Epidemiol* 1988; **41**: 1065– 1074.
3. Bassett SS, Magaziner J., Hebel JR. Reliability of proxy response on mental health indices for aged, community dwelling women. *Psychol Aging* 1990; **5**: 127– 132.
4. Magaziner J., Zimmerman SI, Gruber-Baldini AL *et al*. Proxy reporting in five areas of functional status: Comparison with self-reports and observations of performance. *Am J Epidemiol* 1997; **146**: 418– 428.

5. Weinberger M., Samsa GP, Schmader K. *et al*. Comparing proxy and patients' perceptions of patients' functional status: Results from an outpatient geriatric clinic. *J Am Geriatr Soc* 1992; **40**: 585– 588.

6. Magaziner J.. The use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer's Disease and related outcomes. *Alzheimer Dis Assoc Disord* 1997; **11**(Suppl 6): 168– 174.

7. Moye J., Robiner WN, Mackenzie TB. Depression in Alzheimer patients: Discrepancies between patient and caregiver reports. *Alzheimer Dis Assoc Disord* 1993; **4**: 187– 201.

8. Logsdon RG, Teri L.. Depression in Alzheimer's disease patients: Caregivers as surrogate reporters. *J Am Geriatr Soc* 1995: 150– 155.

9. Kiyak HA, Teri L., Borson S.. Physical and functional health assessment in normal aging and in Alzheimer's disease: Self-reports vs family reports. *Gerontologisi* 1994; **34**: 324– 330.

10. Hays RD, Vickrey BG, Hermann BP *et al*. Agreement between self reports and proxy reports of quality of life in epilepsy patients. *Qual Life Res* 1995; **4**: 159– 168.

11. Graham P., Jackson R.. Primary versus proxy respondents: Comparability of questionnaire data on alcohol consumption. *Am J Epidemiol* 1993; **138**: 443– 452.

12. Corder LS, Woodbury MA, Manton KG. Proxy response patterns among the aged: Effects on estimates of health status and medical care utilization from the 1992–1984 Long-Term Care Surveys. *J Clin Epidemiol* 1996; **49**: 173– 182.

13. Grootendorst PV, Feeny DH, Furlong W.. Does it matter whom and how you ask? Inter- and intra-rater agreement in the Ontario Health Survey. *J Clin Epidemiol* 1997; **50**: 127– 135.

14. Halabi S., Zurayk H., Awaida R., Darwish M., Saab B.. Reliability and validity of self and proxy reporting of morbidity data: A case study from Beirut, Lebanon. *Int J Epidemiol* 1992; **21**: 607– 612.

15. Teri L., Truax P.. Assessment of depression in dementia patients: Association of caregiver mood with depression ratings. *Gerontologist* 1994; **34**: 231– 234.

16. Duke University Center for the Study of Aging and Human Development. *Multidimensional functional assessment: The OARS methodology*. Durham, NC: Duke University, 1978.

17. Zarit SH, Witlatch CJ. Institutional placement: Phases of the transition. *Gerontologist* 1992; **32**: 665– 672.

18. Miller B., McFall S.. The effect of caregiver's burden on change in frail older persons' use of formal helpers. *J Health Soc Behav* 1991; **32**: 165– 179.

19. Stewart AL, Hays RD, Ware JE. The MOS short-form general health survey. Reliability and validity in a patient population. *Med Care* 1988; **26**: 724– 735.

20. Landis JR, Koch GG. Agreement measures for categorical data. *Biometrics* 1977; **12**: 153– 157.

21. Kanten DN, Mulrow CD, Gerety MB *et al*. Falls: An examination of three reporting methods in nursing homes. *J Am Geriatr Soc* 1993; **41**: 662– 666.

22. Lachapelle D., Dcsaulniers G., Bujold N.. Analysis of repeated measures experiments on individuals subjected to different treatments. *Community Dent Oral Epidemiol* 1988; **16**: 95– 97.

23. Tabachnick B., Fidell LS. *Using multivariate statistics*, 22nd Ed. California State University (Northridge): Harper Collins Publishers, 1989.

24. Hadjistavropoulos T., Taylor S., Tuokko H., Beattie BL. Neuropsychological deficits, caregivers' perception of deficits and caregiver burden. *J Am Geriatr Soc* 1994; **42**: 308– 314.

25. Thompson EH Jr, Futterman AM, Gallagher-Thompson D. *et al*. Social support and caregiving burden in family caregivers of frail elders. *J Gerontol* 1993; **48**: S245– S254.

26. Burns C., Archbold P., Stewart B., Shelton K.. New diagnosis: Caregiver role strain. *Nurs Diagn* 1993; **4**: 70– 74.

27. McFall S., Miller BH. Caregiver burden and nursing home admission of frail elderly persons. *J Gerontol* 1992; **47**: S73– S79.

28. Lawton MP, Moss M., Kleban MH *et al*. A two-factor model of caregiving appraisal and psychological well-being. *J Gerontol* 1991; **46**: P181– 189.

29. Baillie V., Norbeck JS, Barnes LEA. Stress, social support, and psychological distress of family caregivers of the elderly. *Nurs Res* 1988; **37**: 217– 222.

30. Fleming BE, Pendergast DR. Physical condition, activity pattern, and environment as factors in falls by adult care facility residents. *Arch Phys Med Rehabil* 1993; **74**: 627– 630.

31. Magaziner J.. The use of proxy respondents in health studies of the aged. In: RB Wallace, RF Woolson, eds. *The Epidemiologic Study of the Elderly*. New York: Oxford University Press, 1992, pp 120– 129.